人脸识别：回顾与展望

王超名　陈旭东
北京交通大学　北京　100044
oujago@gmail.com

摘要： 人脸识别是人工智能在智能感知领域的一项重要任务，同时也具备巨大的实用价值。本文回顾人脸识别技术在过去几十年来的发展历程及主要成就，并对新近发展起来的基于深度学习的人脸识别方法进行阐述和讨论，最后对深度学习人脸识别方法的未来研究方向做出展望。

关键词： 人脸识别, 人工智能

Face Recognition：Retrospect and Prospect

Chao-Ming Wang　Xu-Dong Chen

Beijing Jiaotong University　Beijing　100044

Abstract: Face recognition is an important perception technology of artificial intelligence. Meanwhile, it has great practical value. This paper recalls the developments and achievements of face recognition technology over the past few decades, summarizes the latest progress of deep learning based methods and points out the possible future directions of deep learning based methods.

Keywords: Face Recognition, Artificial Intelligence

1　引言

人工智能是人类长期以来一直追求的目标。如何创造出智能的机器（intelligent machines），以期其能拥有知识（knowledge），具备推理（reasoning）、计划（planning）、学习（learning）、感知（perception）、交流（communication）、以及移动和操纵物体的能力 [1, 2, 3, 4]，是人工智能一直以来不懈努力的方向。人脸识别 [5] 作为人工智能的一项重要任务，是人工智能技术在智能感知方向上的一个重要领域。人脸识别也具备巨大的实用价值。传统个人身份鉴别使用 ID 卡和密码等手段，但是极易模仿、复制、盗窃，系统无法区分实际使用者，难以区分真正的用户。在这样的背景下，人脸识别作为一种有效的生物特征识别技术，为真正可靠的身份鉴定带来了可能性 [6]。也因此，人脸识别几十年来一直受到众多研究学者的关注，并被广泛应用于如视频监控、访问控制等众多信息安全领域。

一个完整的人脸识别系统主要包括五个组成部分，如图 1，分别为：人脸图像采集、人脸检测、人脸图像预处理、人脸图像特征提取以及匹配与识别。但是，建造一个高性能、高鲁棒性的自动人脸识别系统却是一个极其复杂和困难的事情。实际应用中，光照、对比度、抖动、焦点、模糊、遮挡、分辨率、姿态、表情、噪声等因素 [9]，都会引起人脸面部特征发生很大改变。因此几十年来，为了在各种复杂情况下也能得到人脸图像的最优描述，研究者们提出了一系列的算法与理论。

人脸识别算法大致经历了四个阶段。第一个阶段 (1964 年 – 1991 年) 属于人脸识别的起步探索阶段，人们使用一些简单的算法来初步尝试人脸的机器自动识别。第二阶段 (1991 年 – 1998 年) 是人脸识别的快速发展时期，在条件严格控制的人脸识别任务上人们取得了一些初步的成果，也涌现出了一些对后一阶段极具影响力的算法与理论。第三阶段 (1998 年 – 2014 年) 的人脸识别在上一阶段的基础上，针对姿态、光照、表情、遮挡等外界影响因素，提出了一系列的改进算法与新的理论。这一阶段，人脸识别开始逐渐成熟，一些实用的系统开始诞生。然而，前三个阶段的人脸识别算法实质上大多数都是一种浅层学习 (shallow learning) 模型 [10]，这些模型 (如 SVM, Boosting 等) 的结构基本上可以看成一层隐层节点。尽管浅层模型在理论和应用中都获得了巨大成功，却存在着一定的局限性。如，它们表示复杂函数的能力有限，面对复杂的分类问题时模型的泛化能力受到了一定的制约 [11]。因此在第四阶段 (2014 年 – 至今)，人脸识别的主流算法开始转为深度学习 [12]，深度学习的典型代表便是含有多层隐含节点的深度神经网络 (deep neural

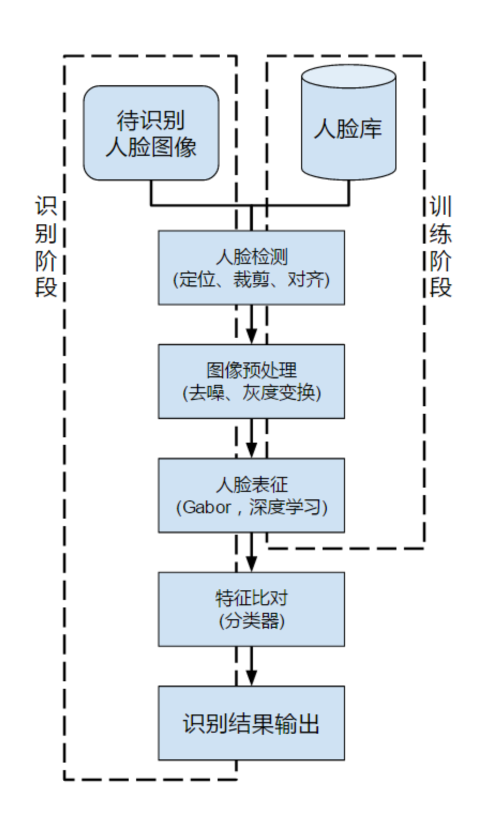


图 1: 人脸识别系统的组成。首先，通过摄像镜头等装置获取到人脸图像，然后基于 Adaboost [7] 或 Cascaded CNN [8] 等算法检测到人脸，挑出有用信息。之后，光线补偿、灰度变换、直方图均衡化等方法用于人脸图像预处理。接着，构建人脸特征描述模型，抽取人脸图像的有效特征。最后，将提取到的特征与数据库中存储的特征模板进行搜索匹配。

networks, DNN)，而大数据、大模型、大计算则是深度神经网络的三大支柱。第四阶段是人脸识别的高潮期，大量实用的系统与成功的应用案例出现，许多新兴的人脸识别公司也开始诞生，人脸识别进入了一个新的时代。

本文将以人脸识别面临的挑战和算法为主线，回顾人脸识别技术在过去几十年来的发展历程及主要成就，并对新近发展起来的深度学习人脸识别方法进行阐述和讨论，最后对基于深度学习的人脸识别的未来研究方向做出展望。

2 第一阶段（1964 年 – 1991 年）

关于人脸识别技术的研究可以追溯到 Francis Galton 分别于 1888 年 [13] 和 1910 年 [14] 在 *Nature* 杂志上发表的有关如何利用人脸图像进行身份认证的文章。他从认知心理学的角度分析了人类自身识别人脸的能力。但真正意义上的人脸识别的工作起源于 Bledsoe 和 Chan 于 1965 年 [15] 在 *Panoramic Research Inc* 上发表的有关人脸自动识别的技术报告。自此，人脸识别技术开启了第一阶段的研究。这一阶段的人脸识别被当作一个一般性的模式识别问题来研究 [16]，所采用的技术主要围绕人脸面部器官之间 (如眼镜、鼻子、下巴等) 的几何结构特征进行展开 [17, 18, 19]，因此这一技术也被称为基于几何特征 (geometric feature based) 的方法。

1966 年，美国德克萨斯大学的 Bledsoe [20, 21] 利用几何特征方法，研制出了第一个半自动的人脸识别系统。该系统首先手动选择眼睛、鼻子、嘴部以及下颌等特征点，然后根据这些特征点计算出距离或角度参数值，如两眼瞳孔之间的距离、两眼瞳孔与鼻尖点的角度、鼻尖与两嘴角间的距离等。同时，为了能够比较不同尺度下的人脸图像，该系统还将所得到的特征向量进行了标准化的处理。其后，贝尔实验室的 Harmon、Goldstein 等人 [22] 开发出了一个基于特征比较的交互式人脸识别系统。该系统使用 21 个特征值来构建人脸识别的参数向量。虽然其识别效果较好，但特征点的选择还是人工进行的。1973 年，卡耐基梅隆大学的 Kanade [23] 提出了基于距离比例的自动特征提取方法，开发出了第一个自动的人脸识别系统。该人脸图像识别系统使用投影法来确定人脸图像的眼睛、鼻子、嘴部等局部特征，通过计算不同特征点之间组成的距离、角度、面积等参数值来得到人脸的特征向量，从而用于人脸图像的比较与识别。然而，上述基于几

何特征的方法对人脸图像有严格的要求，比如人脸图像必须为正面人脸图像，也不能出现形变或旋转。针对这些缺陷，哈佛大学 Yuille 等人 [24, 25] 基于 Fischler 和 Elschlager [26] 的工作，提出了使用可变化的参数模型来表示人脸特征的方法。其中，每个特征都对应于一个参数模型，每个模型构造相应的能量函数，最后利用梯度下降法来寻找能量函数的最优值，进而查找出各部分的人脸特征。

基于几何特征的人脸识别方法的优点在于，描述人脸图像的特征向量十分简洁，物理意义明确，易于理解和应用，对光照变化不敏感，识别速度快。但是由于这种算法过于简单，仅仅利用了面部的结构信息，忽略了局部的细微特征以及纹理信息，因此造成了所需信息的丢失，是一种比较粗犷的人脸表达。而且，人脸二维图像几何特征以及特征点的定位其实很不容易，往往出现由于定位不精确而导致特征点较大偏移的现象。总体说来，这种方法的人脸识别精度不高。

3 第二阶段（1991 年 – 1998 年）

第二阶段人脸识别快速发展，出现了一系列经典的理论与算法，是人脸识别技术的快速发展期 [16]。

1991 年，美国麻省理工学院的 Turk、Pentland 及其他学者 [27, 28] 将基于统计学的主成分分析 (principal component analysis, PCA) 方法引入到人脸识别中，提出了著名的 Eigenface 方法。PCA 也被称作 KL 变换 (Karhunen and Loeve transformation) [29]，其原理简单，容易编程，速度很快，并且识别效果好，可以解决一定的实际问题。至今，仍然有很多人脸识别方法将 PCA 作为特征提取的一个重要预处理步骤。然而 PCA 方法也存在着一些不足。1、PCA 本质上依赖于训练图像和测试图像的灰度相关性，所以算法对人脸图像的亮度、偏移、背景和姿态变化的适应性较差；2、PCA 对小样本的特征提取效果还可以，但对大样本的提取性能却并不理想。

为了克服 PCA 的缺陷，1996 年 Belhumeur 等 [30, 31] 结合 PCA 与线性判别分析 (linear discriminant analysis, LDA) 提出了著名的 Fisherface 方法。不同于无监督的 PCA，LDA [30] 通过考虑数据自身携带的类别标签信息进行特征提取，是一种有监督的方法。LDA 的目标是，寻找一组投影向量，使得数据投影到低维空间后，具有最大类间散列度 (between-class scatter) 和最小的类内散列度 (within-class scatter)。正因为考虑到了高维数据所携带的标签信息，LDA 往往能优化图像数据的低维表示 [32, 33]，更好应对光照、姿态等问题。但 LDA 方法仍然存在着不足，主要表现在以下几个方面。1，LDA 方法要求所观察到的高维数据必须符合高斯分布，然而人脸数据未必符合高斯分布；2，由于数据类间散列度矩阵中的非零特征值最多只能有 $C-1$ 个 (C 为训练样本的类别数)，LDA 方法只能将原始数据最多降到 $C-1$ 维，也因此 LDA 不能直接使用，往往结合 PCA 进行特征提取；3，小样本问题 (small sample size, SSS)：在实际应用领域中，训练样本的特征维度 D 远大于训练样本的数量 N，即 $D \gg N$，从而容易导致类内散列度为奇异矩阵。然而毫无疑问，PCA 和 LDA 是当时重要的理论成果。随后大量沿着 PCA 和 LDA 的思想路径的新方法被提出, 比如基于非负矩阵分解 (non-negative natrix factorization, NMF) [34, 35]、基于核 (kernel) 方法 [36, 37, 38] 等的多种子空间人脸识别算法。

在此阶段，还涌现了其他一些重要的理论与技术。1993 年，Brunelli 和 Poggio[39] 结合实验对比分析了基于结构特征的人脸识别方法与基于模板匹配的人脸识别方法的性能，提出基于模板匹配的方法要优于基于结构特征的方法。这一结论很大程度上促进了基于表观（appearance-based）的线性子空间建模和基于统计模式识别技术的人脸识别方法的发展 [16]。1996 年，洛克菲勒大学的 Penev 和 Atick [40] 提出了局部特征分析 (local feature analysis, LFA) 方法，该方法首先利用 PCA 建立一组局部的特征向量，然后利用稀疏 (sparsification) 技术来得到一组相关性最少且附加有拓扑索引的特征集合，最后选择一组核函数来表征人脸的局部特征。LFA 方法的优点是，它不仅能用低维的数据表示人脸空间，而且克服了 PCA 等方法只关注全局而忽略人脸图像拓扑结构以及忽略人脸局部特征 (如眼镜、鼻子、嘴部等) 的问题。

这一阶段，较有影响的弹性图匹配 (elastic graph matching, EGM) [41, 42, 43] 也被提出。EGM 使用二维结构的 Gabor 小波对人脸图像进行处理，将人脸表达成由若干个特征点构成的具有一定拓扑结构信息的人脸弹性图。弹性图的顶点代表面部关键特征点，其属性为相应特征点的多分辨率、多方向的局部特征；弹性图的边则代表不同特征点之间的几何关系。在进行图匹配时，要在待识别的人脸图像上进行点格阵的全局搜索和局部搜索，查找最相似的点格阵的匹配。该方法的优点是既保留了面部的全局结构特征，也对人脸的关键局部特征进行了建模。同时，小波变换受光照、表情、图像尺寸等因素的干扰较少，因此具备一定的鲁棒性。但是，该方法也存在着不足，因为匹配过程需要反复比较，所以计算量较大，识别速度较慢。

另外，柔性形状模型 (flexible appearance models) [44, 45, 46, 47, 48] 作为人脸建模上的重要模型，也在这一阶段被提出。人脸的形状不是一成不变的，同一个人的人脸形状在不同时期、不同姿态下都会有所变化，而固定的人脸模型很难建模、表示出这些人脸变化。因此，Yuille 等人

[44, 45] 提出了使用参数化的可变模型技术来表示人脸各部分的局部特征，通过模板的偏移、旋转或者形变等操作，来得到模板的最佳匹配。而 Lanitis 等人 [46, 47, 48] 基于 Yuille 等人的思想，提出了成熟的柔性形状模型技术。该技术由两个阶段组成：建模阶段和识别阶段。建模阶段首先得到人脸的形状模型和人脸的灰度值分布模型，然后在识别阶段，利用上一阶段得到的形状和灰度值分布，进行身份的识别。

4　第三阶段（1998 年 – 2014 年）

这一阶段是人脸识别技术的成熟期，主要针对人脸识别中姿态、光照、表情、噪声、遮挡等外界因素变化，以及第二阶段涌现的算法的问题，提出了一系列新的算法与理论。

在 Eigenface 的基础上，为了解决光照、表情等的问题，Moghaddam 和 Pentland 等 [49, 50] 提出了概率主成分分析 (probabilistic pricinpal component analysis, PPCA) 方法。PPCA 在计算两幅人脸图像之间的差异时，既考虑了人脸的不同导致图像差异的可能性，也考虑了光照、表情等因素所引起的差异的可能性。因此该方法对人脸表情和光照的变化具有更好的鲁棒性。另外，一些研究者注意到特征值大的特征向量可能并不是特征提取最好的方向。Cappelli 等 [51] 于是提出了多空间 KL 变换，把训练集分割成多个不同的子集，这些子集采用不同的 KL 变换构造不同的子空间来表示不同的子模式。另外，传统 PCA 方法是一种线性方法，难以发现高维非线性结构数据的内在结构。Kim 等 [36] 据此提出了核主成分分析 (kernel principal component analysis, KPCA) 方法，通过非线性变换转换人脸图像的特征空间，以此挖掘高维数据内在的非线性结构。随后，另外一些 PCA 的扩展方法也相继提出。Vasilescu 和 Terzopoulos [52] 提出了多线性子空间 (multi-linear space, MLS) 方法，在一种多线性框架下进行子空间分析，利用张量分解算法 N-node SVD，在多个相互关联的特征空间上进行维数约简。Yang 和 Zhang [53] 提出了二维 PCA (two-dimensional PCA) 的人脸识别方法，该方法首先用一种二维图像矩阵表示人脸，然后根据二维图像矩阵构建出协方差矩阵，采用该协方差矩阵的主要特征向量作为表示特征，最后进行人脸识别。Cavalcanti 等 [54] 提出了 Eigenbands 的人脸识别算法，首先将人脸图像分解成水平和垂直的条带，然后采用 PCA 方法为每一条带抽取特征信息。甚至在近几年，PCA 方法仍然持续有人研究。Kadam [29] 将 PCA 和离散余弦转换 (discrete cosine transform, DCT) 结合起来用于人脸数据降维, 实验结果显示这种混合方法保证识别速度的同时，能获得比简单 PCA 更高的准确率。Bakhshi 等人 [55] 先使用 SIFT(scale invariant feature transform) [56] 和 SURF(speeded sp robust features) [57] 提取人脸图像的特征，之后使用 PCA 处理图像，能在光照、姿态、旋转等条件下获得更高的识别率。Poon 等 [58] 通过实验检验了各种不同的光照不变技术 (illumination invariant techniques)，发现其中一种 Gradientfaces 的技术在数据预处理阶段结合 PCA，能显著提升人脸识别准确率。Barnouti 和 N.H. [59] 提出了一种 BP 神经网络、PCA 和 DCT 的混合方法。其中，BP 神经网络结合 PCA 能更容易识别人脸，DCT 能压缩人脸数据并提升识别速度。

在 Fisherface 的基础上，为了解决 LDA 的小样本问题，Chen 等 [60] 提出了一种新的 LDA 算法用于求取类内散列度矩阵零空间中的最优判别矢量 (optimal discriminant vector)，并取得了较好的效果。Wang 和 Tang [61] 采用随机子空间和融合的方法用于改进 Fisherface 和 N-LDA(Nullspace LDA) 算法。他们还提出基于一种概率视觉模型的双重空间 LDA 算法 [62]，用于解决小样本问题。Howland 和 Ye 等人 [63, 64] 采用 GSVD (generalized singular value decomposition) 算法来解决传统 LDA 算法中的散列度矩阵奇异的问题。Lu 等 [65, 66, 67] 将核方法引入到 LDA 中，提出了基于核的 LDA 算法用于解决人脸识别中的人脸模式分布的非线性和小样本问题。Jing 等 [68] 将非相关最优判别矢量和 Fisherface 结合，用于改进传统的 LDA 算法。Yang 等 [69] 提出了一种局部特征判别分析的方法来解决小样本问题。Liu 等 [37, 70, 71] 采用余弦核函数来提高判别分类的能力，同时采用基于几何的特征向量选择机制来减少和判别分析算法的计算复杂度。甚至在最近几年也提出了 LDA 的改进算法。Murtaza 等人 [72] 提出了 AMFC-LDA(adaptive margin fisher's criterion linear discriminant analysis) 算法来克服传统 LDA 和最大间距准则 (maximum margin criterion, MMC) 的问题。AMFC-LDA 算法不再存在小样本问题，不仅拥有较低的错误拒绝率 (false rejection rate) 和错误接受率 (false acceptance rate)，而且计算复杂度也得到一定程度的降低，同时收敛速度更快。

传统的人脸识别方法大多建立在人脸空间线性可分的假设下，但从 2000 年开始，逐渐有研究表明人脸图像可能位于或近似地位于嵌入到高维空间的低维流形上 [73, 74, 75, 76]。传统的线性方法不能表达出人脸空间的凸起与凹进，这成为了人脸识别突破的瓶颈问题。于是，流形学习 (manifold learning) [77, 78] 技术开始被引入到人脸识别中。流形学习的本质是挖掘高维数据的内在规律以及本征结构。流形学习所得到的子空间被称为嵌入空间，它是一种非线性空间，并保持了原始样本空间的全局和局部拓扑结构，接近于人类的视觉感知系统。相比于线性空间，非线性空间对人脸特征具有更好的表达。基于流形学习的人脸识别方法可分为无监督、有监督和半监督三类。表 1 列出了一些基于流形学习的代表算法。

类别	代表算法
无监督	等距离特征映射（Isometric Feature Mapping，ISOMAP）[74]
	局部线性嵌入（Locally Linear Embedding，LLE）[75]
	拉普拉斯特征映射（Laplacian Eigenmaps，LE）[79]
	局部保持投影（Local Preserving Projections，LPP）[80]
	无监督判别投影（Unsupervised Discriminant Projection，UDP ）[81]
	邻域保持嵌入（Neighborhood Preserving Eebedding，NPE）[82]
	正交邻域保持投影（Orthogonal Neighborhood Preserving Projections，ONPP）[83]
	多视角邻域保持投影（Multi-View Neighborhood Preserving Projections，Multi-NPP）[84]
	稀疏保持投影（Sparsity Preserving Projections，SPP）[85]
	图优化局部保持投影（Graph Optimized Locality Preserving Projections，GoLPP）[86]
	稀疏约束的图优化维数约简（Graph Optimization for Dimensionality Reduction with Sparisity Constrints，GODRSC）[87]
	自适应图维数约简（Dimensionality Reduction with Adaptive Graph，DRAG）[88]
有监督	边缘费舍尔分析（Marginal Fisher Analysis，MFA）[89]
	最大边缘准则（Maximum Margin Criterion，MMC）[90]
	局部敏感判别分析（Locality Sensitive Discriminant Analysis，LSDA）[91]
	局部判别嵌入（Local Discriminant Embedding，LDE）[92]
	局部判别投影（Local Discriminant Projections，LDP）[93]
	监督局部保持投影（Supervised Localitiy Preserving Projections，SLPP）[94]
	多流形判别分析（Multi-Manifold Discriminant Analysis，MMDA）[95]
	判别多流形分析（Discriminative MultimanifoldAnalysis，MDA）[96]
	多流形局部线性嵌入（Multiple Manifold Locally Linear Embedding，MM-LLE）[97]
半监督	半监督判别分析（Semi-supervised Discriminant Analysis，SDA ）[98]
	半监督子流形判别分析（Semi-supervised Sub-manifold Discriminant Analysis，S^3MDA）[99]
	半监督局部费舍尔判别分析（Semi-supervised Local Fisher Discriminant Analysis，SELF）[100]
	多流形半监督学习（Multi-manifold Semi-supervised Learning，MMSSL）[101]

表 1: 基于流形学习的代表算法。

尽管上述大量的线性和非线性方法在人脸识别中已取得了较好的性能，但真实情境下它们还是容易受人脸图像的姿态、光照、表情等的变化的影响。本质上来说，上述方法都是基于人脸全局信息的人脸识别方法，得到的人脸描述特征被成为全局特征。全局特征是指其特征向量的每一维都包含了人脸图像上所有部分 (甚至所有像素) 的信息, 因此反映的是人脸的整体属性 [102]。与此相反，另一类人脸识别方法得到的人脸描述特征被称为局部特征。局部特征的每一维都只对应着人脸图像上的一个局部区域, 因此这类方法侧重于提取人脸的细节特征 [102]。局部特征对人脸的光照、表情和遮挡等变化不敏感, 因此被越来越多地应用于人脸表示中。Gottumukkal 等 [103] 提出模块化主成分分析 (modular PCA，ModPCA)，其首先将人脸图像划分为若干个小的子图像或子模式，然后将所有子模式看作整体并利用 PCA 提取子模式的特征，最后所有子模式集的特征整合成全局特征用于人脸的识别。但 ModPCA 忽略了子模式空间结构的位置信息，因此 Chen 和 Zhu [104] 提出了子模块主成分分析 (sub-pattern based PCA, SpPCA)。不同于 ModPCA 将所有子模式看成整体，SpPCA 将原始人脸图像相同位置的所有子图像组成子模式集，并对每个子模式集利用 PCA 提取子特征向量集。后续还提出了自适应加权子模式主成分分析 (adaptively weighted sub-pattern PCA, Aw-SpPCA) [105] 和交叉子模式相关主成分分析 (cross-sub-pattern based PCA, SubXPCA) [106]，用于 SpPCA 方法的改进。当然除了 PCA，另外一些常用的局部特征方法有: 局部二值模式 (local binary pattern, LBP) [107]、Gabor 小波 [108, 109, 110]、局部非负矩阵分解 (local non-negative matrix factorization, LNMF) [111, 112] 等。值得一提的是, 基于 Gabor 小波的人脸识别方法 (比如 Gabor-fisher classifier [108], local Gabor binary pattern [109], weighted sub-Gabor [110]) 在很多公开数据库上和性能评测中取得了非常好的结果, 受到了众多研究者的关注。Gabor 小波也因此被认为是一种非常有效的人脸表示方法。

2006 年，Donoho 等人 [113] 提出了压缩感知 (Compressive Sensing，CS) 理论。2009 年，Wright

等人 [114] 将 CS 理论应用于人脸识别问题中并提出了稀疏表示分类 (sparse representatin based classification，SRC) 的方法。这类方法也被称为稀疏编码 (sparse coding) 方法。随后，大量稀疏编码方法开始被提出。基于稀疏编码的人脸识别方法从另外一个新的角度来看待和处理人脸识别问题，其基本思想为：人脸测试样本可以由若干个训练样本近似表达，且测试样本属于该近似表达中占比最大的类别的概率最大。由于在测试样本和训练样本之间架设了一座桥梁，这类方法往往能取得优异的性能。即使在脸部存在遮挡等复杂情况下，稀疏编码方法仍能取得较好的人脸识别性能。也正因此，该方法被评价为人脸识别领域的重大突破之一。稀疏编码算法大致分为五类 [115]：重构稀疏编码 (reconstructive sparse coding), 有监督稀疏编码 (supervised sparse coding), 判别稀疏编码 (discriminative sparse coding), 结构稀疏编码 (structured sparse coding) 和图正则化稀疏编码 (graph regularized sparse coding)。

- 重构稀疏编码：这类方法设计不同的优化算法来学习最优字典，通过最小化数据重构误差 (data reconstruction error)，来找到相应的稀疏表达系数。代表性算法有匹配追踪 (matching pursuit, MP) [116]、正交匹配追踪 (orthogonal matching pursuit, OMP) [117]、基追踪 (basis pursuit, BP) [118] 等。

- 有监督系数编码：这类方法通过人脸图像高维数据自身携带的标签类别信息来学习一个超完备字典 (over-completed dictionary) 与相应的稀疏表达系数。Plam 等 [119] 提出一种字典重构与分类器学习的联合框架，来考虑类别标签和线性预测分类误差。Zhang 等 [120] 将字典学习与分类器参数学习整合到一个目标函数中，提出了判别 KSVD(discriminative K-SVD, D-KSVD)。Jiang 等 [121] 进一步整合数据的类别标签信息和分类误差，对 D-KSVD 方法进行了扩展，提出了类别标签一致 K 均值奇异分解 (label consistent KSVD, LC-KSVD) 算法。

- 判别稀疏编码：不同于有监督稀疏编码方法直接利用数据的类别标签信息，这类方法将类可分性准则 (class separability criterion) 整合到稀疏编码目标函数中。比较常用的类可分性准则有：softmax 函数 [122]、费舍尔判别准则 (Fisher discrimination criterion) [123]、Hinge 损失函数 [124] 等。

- 结构稀疏编码：这类方法通过人脸图像的先验知识来修改惩罚约束项，促使学到的特征按照一定的规则排列，从而使其学到具有一定结构特性的字典。结构稀疏编码主要利用组稀疏 (group sparse) [125] 与层次稀疏 (hierarchical sparse) [126] 来对重构稀疏编码方法进行扩展。同时，Jia 等 [127] 将结构稀疏编码引入多视角学习 (multi-view learning) 框架中，以期学习到一个潜在的子空间；Zhang 等 [128] 在人脸识别任务中，则将多视角分类问题看成联合稀疏表示问题。

- 图正则化稀疏编码：这类方法主要在稀疏编码过程中采用不同的图正则化来保持数据的局部几何结构关系。Zheng 等 [129] 将图拉普拉斯正则项 (Laplacian regularization, LR) 引入到稀疏编码框架中，用于保持数据分布的局部几何结构。Zheng 等 [130] 基于二阶海森能量 (second-order Hessian energy) 提出了海森稀疏编码 (Hessian sparse coding) 来更好地保持数据的局部拓扑关系。Gao 等 [131] 利用超图 (hypergraph) 相对于传统图模型能更有效地表达存在高阶关系变量的特性，提出超图拉普拉斯正则化稀疏编码 (hypergraph Laplacian sparse coding) 来保持特征空间局部一致性。

稀疏编码方法与常规降维方法实质上具有相似的目标。即，为样本提供一个某种意义上的最优描述，以获得较高分类正确率。但稀疏编码方法不同于常规降维方法的主要之处在于：常规降维方法仅仅依据所有训练样本来为这些训练样本产生最优描述，而据此对测试样本产生的描述结果却不一定最优；然而，稀疏编码方法是同时利用所有训练样本与当前测试样本来为当前测试样本提供一个最优描述。

5 第四阶段（2014 年 – 至今）

在 2014 年之前，人脸识别的主要技术路线是“人造或基于学习的局部描述子 (如 LBP，Gabor) + 度量学习 (distance metric learning, DML)”。但从 2014 年开始，大量基于深度学习的人脸识别方法被相继提出，人脸识别技术的主流技术路线开始转为“深度学习 + 人脸图像大数据”。基于深度学习的人脸识别方法近年来呈现出两种重要的趋势。一是深度学习网络不断变大变深；二是带标注的人脸训练数据不断增多，大数据成为提升人脸识别性能的关键。

2014 年，Facebook 人工智能实验室的 Taigman 等 [132] 提出了 DeepFace 网络，在户外标记人脸数据库 (Labeled Faces in the Wild, LFW) [133] 上取得了 97.25% 的准确率，首次接近人类水平，

可谓是基于深度学习的人脸识别方法的奠基之作。DeepFace 一共八层网络结构；前三层是传统的 CNN 结构，用于提取低层次的特征，比如简单的边和纹理；接下来三层是参数不共享的卷积核，用于提取不同区域不同的统计特征；最后两层是全连接层，作用是捕捉人脸图像不同位置的特征之间的相关性。DeepFace 在训练时使用了 4000 个人的总共 4 百万张图片。另外，与之后的深度学习方法相比，DeepFace 最大的不同在于在训练神经网络前，使用了 3D 对齐方法来解决传统 2D 对齐不能解决的面外旋转 (out-of-plane rotations) 问题。从 2014 年开始, 香港中文大学多媒体实验室孙祎等人先后提出了一系列深度学习人脸识别网络 [134, 135, 136, 137]。与 DeepFace 不同的是，这类网络由一系列相同的小网络融合而成。每个小网络的输入都是人脸图像经过多尺度多通道多区域切分后的一个 patch，之后每个 patch 学到的特征向量通过整合最后得到整体人脸图像的特征向量。在他们最先提出的 DeepID(deep hidden identity features) 网络 [134] 中，每个小网络由 4 个卷积层、3 个 pooling 层和两个全连接层组成，总共 100 个小网络。在 LFW 数据集上取得了 97.45% 的准确率。DeepID2 网络 [135] 沿袭了 DeepID 的基本思路，但在学习特征的时候，该网络不仅考虑了分类准确率，还考虑了类间差距。具体做法就是在目标函数中添加类间差距一项。于是在网络训练时，一要最小化类内变化，二要最大化类间差别。DeepID2 一共 200 个小网络，在 LFW 数据集上取得了 99.15% 的准确率。在 DeepID2+ [136] 中，孙祎等人不仅继续修改了网络结构，还增加了对卷积神经网络的大量的分析，发现卷积神经网络对人脸图像具有适度稀疏性、特征选择性和遮挡鲁棒性等特性。DeepID2+ 由 25 个小网络组成，在 LFW 数据集上取得了 99.47% 的准确率。随后提出的 DeepID3 有两种不同的结构，分别为 DeepID3 net1，DeepID3 net2。相对于 DeepID2+，DeepID3 借鉴了 VGG-Net [138] 的思想，网络层数更多，变得更深。DeepID3 还借鉴了 GoogLeNet [139] 的思路，引入了 Inception 层，在网络中将两个连续的卷积层直接相连，使得整个网络具有更大的感受野 (respective field) 和更复杂的非线性转化，同时还能限制了参数的数量。DeepID3 也由 25 个网络组成，在 LWF 数据集取得了 99.53% 的成绩。另外，为了使得卷积神经网络训练得更加充分，这一系列网络都在训练时通过使用外部数据集 CelebFaces+ 来加大训练数据。

2015 年, 谷歌公司的 Schroff 等 [140] 提出了 FaceNet 算法。该算法利用三元组损失函数进行网络训练以直接将人脸图像映射到欧几里得空间，空间的距离代表了人脸图像的相似性。只要该映射空间生成，人脸识别、验证和聚类等任务都可以一起轻松完成。FaceNet 使用了 2 亿张人脸图像进行训练，在 LFW 测试集上, 该算法取得了 99.63% 的精度。同年，百度公司 Liu 等人 [141] 提出了一种两步学习方法，首先利用 multi-path 深度 CNN 网络在人脸不同区域进行特征提取，然后利用深度度量学习 (deep metric learning) 将前一阶段学到的特征向量降到 128 维。该算法在 18000 人的 120 万人脸数据上进行训练，取得了 LFW 数据集 99.77% 的准确率。另外，腾讯公司 [142]、旷视公司 [143] 等也都提出了自己的基于深度学习的算法。

上述这些成果，几乎宣告了 LFW 数据集从 2008 年到 2015 年以来长达八年的性能竞赛的结束。LFW 数据库是真实条件下的人脸识别问题的测试基准，但对于实际应用中的光照、对比度、抖动、焦点、模糊、遮挡、分辨率、姿态等影响人脸识别的复杂因素 [9] 依然没有得到很好的覆盖。因此，近年来一些更具有挑战性的人脸数据库开始发布，如 IJB-A 数据库 [144]、MegaFace 数据库 [145] 和微软百万名人数据库 [146] 等。近年来的人脸识别的研究也开始逐渐聚焦于这些更具挑战性的实际应用场景。2016 年，南加州大学 Iacopo Masi 等人 [147] 提出一种解决人脸识别中大姿态变化问题的方法。不同于当前其他大部分利用单一模型通过训练大量数据或通过人脸矫正来学习姿态不变性的方法，Iacopo Masi 等人通过使用五个指定角度模型和渲染人脸图片的方法来处理大姿态变化。该方法在 IJB-A 数据库进行评测，取得了不错的效果。中科院计算所 Meina Kan 等人 [148] 针对解决人脸识别中的跨视图或跨姿态问题也提出对应的解决办法。他们尝试移除人脸数据之间的跨模态差异性，并且寻找跨模态之间的非线性的差异性和模态不变性表达。

总体说来，这一阶段是人脸识别的高潮期，人脸识别因为深度学习开始进入一个新的时代。

6 总结与展望

作为生物特征识别的典范，人脸识别经过几十年的发展，已经逐渐变得实用，基于人脸识别的商业公司也层出不穷。本文回顾了人脸识别在过去五十多年的发展历史，介绍了人脸识别在各阶段所取得的成果和面临的挑战。特别是近年来深度学习使得人脸识别得准确率达到了新的高度。但未来基于深度学习的人脸识别的研究还有许多亟须待解决的问题与挑战：

（1）从无标注的数据里学习

大数据已经成为基于深度学习的人脸识别的标配，上百万的人工标注的数据已经成为人脸识别性能提升的必要条件。然而，标注数据的获取却是十分困难、十分昂贵的。因此，当前深度学习的一个前沿研究热点就是如何从无标注的数据里进行学习。目前，生成式对抗网络 (Generative

Adversarial Nets, GANs) [149] 与对偶学习 (dual learning) [150] 有望在此问题上获得一些突破。

（2）降低模型的大小

目前，基于深度学习的人脸识别模型大小一般在 500M 以上，甚至是上 G 的大小。这样的模型难以在手机等移动设备上使用。因此，如何把大模型变成小模型，降低设备内存或存储空间的消耗，以及降低设备的能耗，是当前基于深度学习的人脸识别的一个努力的方向。

（3）更快速有效的训练方法

大数据与大模型是深度学习人脸识别方法的支柱，但大数据与大模型带来的一个问题是大计算。算法训练的计算量大大增加，往往需要几块甚至上十块最先进的 GPU 学习训练几周甚至上月的时间。这对于模型的调参来说是一个极其困难的事情。因此，如何设计更高级、更快速、更有效的算法，是基于深度学习的人脸识别算法的又一挑战。

（4）小样本学习

大数据是基于深度学习的人脸识别方法的性能提升的关键，但是，基于大数据的学习方式与人类的智能很不一样。人往往能从小样本进行学习，具备举一反三、领域迁移的能力。人对于一个人的人脸的识别，往往只需要几张照片就能准确分类。原因在于，人经过有限的训练，结合规则、知识、经验，能够应付各种复杂的情况。但是当前基于深度学习的人脸识别方法并不具备逻辑思考、联想和推理等能力，必须依靠大数据来覆盖各种可能的情况。因此，正如张钹教授 [151] 所说，知识驱动与数据驱动的结合，或许是人工智能的突破点之一。

参考文献

[1] Stuart Russell, Peter Norvig, and Artificial Intelligence. A modern approach. Artificial Intelligence. Prentice-Hall, Egnlewood Cliffs, 25:27, 1995.

[2] Nils J Nilsson. Artificial intelligence: a new synthesis. Elsevier, 1998.

[3] David Lynton Poole, Alan K Mackworth, and Randy Goebel. Computational intelligence: a logical approach, volume 1. Oxford University Press New York, 1998.

[4] George F Luger. Artificial intelligence: structures and strategies for complex problem solving. Pearson education, 2005.

[5] Anil K Jain and Stan Z Li. Handbook of face recognition. Springer, 2011.

[6] Salil Prabhakar, Josef Kittler, Davide Maltoni, Lawrence O'Gorman, and Tieniu Tan. Introduction to the special issue on biometrics: Progress and directions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(4):513–516, 2007.

[7] Paul Viola and Michael J Jones. Robust real-time face detection. International journal of computer vision, 57(2):137–154, 2004.

[8] Hongwei Qin, Junjie Yan, Xiu Li, and Xiaolin Hu. Joint training of cascaded cnn for face detection. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[9] Ayman Abaza, Mary Ann F. Harrison, and Thirimachos Bourlai. Quality metrics for practical face recognition. In ICPR, 2012.

[10] 郭丽丽 and 丁世飞. 深度学习研究进展. 计算机科学, 42(5):28–33, 2015.

[11] Yoshua Bengio et al. Learning deep architectures for ai. Foundations and trends® in Machine Learning, 2(1):1–127, 2009.

[12] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. Nature, 521(7553):436–444, 2015.

[13] Francis Galton. Personal identification and description. Journal of Anthropological Institute of Great Britain and Ireland, pages 177–191, 1889.

[14] Francis Galton. Numeralised profiles for classification and recognition. Nature, 83:127–130, 1910.

[15] Woodrow W Bledsoe and Helen Chan. A man-machine facial recognition system—some preliminary results. Panoramic Research, Inc, Palo Alto, California., Technical Report PRI A, 19:1965, 1965.

[16] 山世光. 人脸识别中若干关键问题的研究. 博士学位论文. 北京: 中国科学院计算技术研究所, 2004.

[17] Michael David Kelly. Visual identification of people by computer. Technical report, DTIC Document, 1970.

[18] Leon D Harmon and Willard F Hunt. Automatic recognition of human face profiles. Computer Graphics and Image Processing, 6(2):135–156, 1977.

[19] LD Harmon, MK Khan, Richard Lasch, and PF Ramig. Machine identification of human faces. Pattern Recognition, 13(2):97–110, 1981.

[20] Woodrow W Bledsoe. The model method in facial recognition. Panoramic Research Inc., Palo Alto, CA, Rep. PR1, 15:47, 1966.

[21] Woodrow Wilson Bledsoe. Man-machine facial recognition. Rep. PRi, 22, 1966.

[22] A Jay Goldstein, Leon D Harmon, and Ann B Lesk. Identification of human faces. Proceedings of the IEEE, 59(5):748–760, 1971.

[23] Takeo Kanade. Picture processing system by computer complex and recognition of human faces. Doctoral dissertation, Kyoto University, 3952:83–97, 1973.

[24] Alan L Yuille, Peter W Hallinan, and David S Cohen. Feature extraction from faces using deformable templates. International journal of computer vision, 8(2):99–111, 1992.

[25] Alan L Yuille. Deformable templates for face recognition. Journal of Cognitive Neuroscience, 3(1):59–70, 1991.

[26] Martin A Fischler and Robert A Elschlager. The representation and matching of pictorial structures. IEEE Transactions on computers, 100(1):67–92, 1973.

[27] Ian Jolliffe. Principal component analysis. Wiley Online Library, 2002.

[28] Matthew Turk and Alex Pentland. Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1):71–86, 1991.

[29] Kiran D Kadam. Face recognition using principal component analysis with dct. International Journal of Engineering Research and General Science, ISSN, pages 2091–2730, 2014.

[30] Anil K Jain, Robert P. W. Duin, and Jianchang Mao. Statistical pattern recognition: A review. IEEE Transactions on pattern analysis and machine intelligence, 22(1):4–37, 2000.

[31] Peter N. Belhumeur, João P Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on pattern analysis and machine intelligence, 19(7):711–720, 1997.

[32] Suman Kumar Bhattacharyya and Kumar Rahul. Face recognition by linear discriminant analysis. International Journal of Communication Network Security, 2(2):31–35, 2013.

[33] Nawaf Hazim Barnouti, Sinan Sameer Mahmood Al-Dabbagh, Wael Esam Matti, and Mustafa Abdul Sahib Naser. Face detection and recognition using viola-jones with pca-lda and square euclidean distance. International Journal of Advanced Computer Science and Applications (IJACSA), 7(5), 2016.

[34] Daniel D Lee and H Sebastian Seung. Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755):788–791, 1999.

[35] Yuan Wang, Yunde Jia, Changbo Hu, and Matthew Turk. Non-negative matrix factorization framework for face recognition. International Journal of Pattern Recognition and Artificial Intelligence, 19(04):495–511, 2005.

[36] Kwang In Kim, Keechul Jung, and Hang Joon Kim. Face recognition using kernel principal component analysis. IEEE signal processing letters, 9(2):40–42, 2002.

[37] Qingshan Liu, Hanqing Lu, and Songde Ma. Improving kernel fisher discriminant analysis for face recognition. IEEE transactions on circuits and systems for video technology, 14(1):42–49, 2004.

[38] Francis R Bach and Michael I Jordan. Kernel independent component analysis. Journal of machine learning research, 3(Jul):1–48, 2002.

[39] Roberto Brunelli and Tomaso Poggio. Face recognition: Features versus templates. IEEE transactions on pattern analysis and machine intelligence, 15(10):1042–1052, 1993.

[40] Penio S Penev and Joseph J Atick. Local feature analysis: A general statistical theory for object representation. Network: computation in neural systems, 7(3):477–500, 1996.

[41] Joachim Buhmann, Martin Lades, and Christoph von der Malsburg. Size and distortion invariant object recognition by hierarchical graph matching. In Neural Networks, 1990., 1990 IJCNN International Joint Conference on, pages 411–416. IEEE, 1990.

[42] Martin Lades, Jan C Vorbruggen, Joachim Buhmann, Jörg Lange, Christoph von der Malsburg, Rolf P Wurtz, and Wolfgang Konen. Distortion invariant object recognition in the dynamic link architecture. IEEE Transactions on computers, 42(3):300–311, 1993.

[43] Laurenz Wiskott, Norbert Krüger, N Kuiger, and Christoph Von Der Malsburg. Face recognition by elastic bunch graph matching. IEEE Transactions on pattern analysis and machine intelligence, 19(7):775–779, 1997.

[44] A L Yuille. Deformable templates for face recognition. Journal of cognitive neuroscience, 3 1:59–70, 1991.

[45] Alan L. Yuille, Peter W. Hallinan, and David S. Cohen. Feature extraction from faces using deformable templates. International Journal of Computer Vision, 8:99–111, 1992.

[46] A Lanitis, CJ Taylor, and TF Cootes. Automatic tracking, coding and reconstruction of human faces, using flexible appearance models. Electronics Letters, 30(19):1587–1588, 1994.

[47] Andreas Lanitis, Christopher J Taylor, and Timothy F Cootes. Automatic face identification system using flexible appearance models. Image and vision computing, 13(5):393–401, 1995.

[48] Andreas Lanitis, Christopher J. Taylor, and Timothy F. Cootes. Automatic interpretation and coding of face images using flexible models. IEEE Transactions on Pattern Analysis and machine intelligence, 19(7):743–756, 1997.

[49] Baback Moghaddam and Alex Pentland. Probabilistic visual learning for object representation. IEEE Transactions on pattern analysis and machine intelligence, 19(7):696–710, 1997.

[50] Baback Moghaddam, Tony Jebara, and Alex Pentland. Bayesian face recognition. Pattern Recognition, 33(11):1771–1782, 2000.

[51] Raffaele Cappelli and Davide Maltoni. Multispace kl for pattern representation and classification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(9):977–996, 2001.

[52] M Alex O Vasilescu and Demetri Terzopoulos. Multilinear subspace analysis of image ensembles. In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, volume 2, pages II–93. IEEE, 2003.

[53] Jian Yang, David Zhang, Alejandro F Frangi, and Jing-yu Yang. Two-dimensional pca: a new approach to appearance-based face representation and recognition. IEEE transactions on pattern analysis and machine intelligence, 26(1):131–137, 2004.

[54] George DC Cavalcanti et al. Eigenbands fusion for frontal face recognition. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, volume 1, pages I–665. IEEE, 2003.

[55] Yukti Bakhshi, Sukhvir Kaur, and Prince Verma. An improvement in face recognition for invariant faces. International Journal of Current Engineering and Technology, 6(2):423–426, 2016.

[56] David G Lowe. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2):91–110, 2004.

[57] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In European conference on computer vision, pages 404–417. Springer, 2006.

[58] Bruce Poon, M Ashraful Amin, and Hong Yan. Improved methods on pca based human face recognition for distorted images. In Proceedings of the International MultiConference of Engineers and Computer Scientists, volume 1, 2016.

[59] Nawaf Hazim Barnouti. Face recognition using pca-bpnn with dct implemented on face94 and grimace databases. METHODOLOGY, 2:0.

[60] Li-Fen Chen, Hong-Yuan Mark Liao, Ming-Tat Ko, Ja-Chen Lin, and Gwo-Jong Yu. A new lda-based face recognition system which can solve the small sample size problem. Pattern Recognition, 33:1713–1726, 2000.

[61] Xiaogang Wang and Xiaoou Tang. Random sampling lda for face recognition. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., 2004.

[62] Xiaogang Wang and Xiaoou Tang. Dual-space linear discriminant analysis for face recognition. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., 2004.

[63] Peg Howland and Haesun Park. Generalizing discriminant analysis using the generalized singular value decomposition. IEEE transactions on pattern analysis and machine intelligence, 26(8):995–1006, 2004.

[64] Jieping Ye, Ravi Janardan, Cheong Hee Park, and Haesun Park. An optimization criterion for generalized discriminant analysis on undersampled problems. IEEE Trans. Pattern Anal. Mach. Intell., 26:982–994, 2004.

[65] Juwei Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. Face recognition using kernel direct discriminant analysis algorithms. IEEE Trans. Neural Networks, 14:117–126, 2003.

[66] Juwei Lu, Konstantinos N Plataniotis, and Anastasios N Venetsanopoulos. Regularized d-lda for face recognition. In Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on, volume 3, pages III–125. IEEE, 2003.

[67] Juwei Lu, Konstantinos N. Plataniotis, and Anastasios N. Venetsanopoulos. Face recognition using lda-based algorithms. IEEE Trans. Neural Networks, 14:195–200, 2003.

[68] Xiao-Yuan Jing, David Zhang, and Yuan Yan Tang. An improved lda approach. IEEE Trans. Systems, Man, and Cybernetics, Part B, 34:1942–1951, 2004.

[69] Qiong Yang, Xiaoqing Ding, and Z Chen. Discriminant local feature analysis of facial images. In Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on, volume 2, pages II–863. IEEE, 2003.

[70] Wei Liu, Yunhong Wang, Stan Z. Li, and Tieniu Tan. Null space-based kernel fisher discriminant analysis for face recognition. In FGR, 2004.

[71] Qingshan Liu, Xiaoou Tang, Hanqing Lu, and Songde Ma. Kernel scatter-difference based discriminant analysis for face recognition. In ICPR, 2004.

[72] Marryam Murtaza, Muhammad Sharif, Mudassar Raza, and J Shah. Face recognition using adaptive margin fisher’s criterion and linear discriminant analysis. International Arab Journal of Information Technology, 11(2):1–11, 2014.

[73] H Sebastian Seung and Daniel D Lee. The manifold ways of perception. science, 290(5500):2268–2269, 2000.

[74] Joshua B Tenenbaum, Vin De Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. science, 290(5500):2319–2323, 2000.

[75] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. science, 290(5500):2323–2326, 2000.

[76] Ya Chang, Changbo Hu, and Matthew Turk. Manifold of facial expression. In AMFG, 2003.

[77] J B Tenenbaum, V de Silva, and J C Langford. A global geometric framework for nonlinear dimensionality reduction. Science, 290 5500:2319–23, 2000.

[78] Lawrence Cayton. Algorithms for manifold learning. Univ. of California at San Diego Tech. Rep, pages 1–17, 2005.

[79] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In NIPS, 2001.

[80] Xiaofei He and Partha Niyogi. Locality preserving projections. In NIPS, 2003.

[81] Jian Yang, David Zhang, Jing-Yu Yang, and Ben Niu. Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics. IEEE Trans. Pattern Anal. Mach. Intell., 29:650–664, 2007.

[82] Xiaofei He, Deng Cai, Shuicheng Yan, and HongJiang Zhang. Neighborhood preserving embedding. In Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1, 2005.

[83] Effrosini Kokiopoulou and Yousef Saad. Orthogonal neighborhood preserving projections: A projection-based dimensionality reduction technique. IEEE Trans. Pattern Anal. Mach. Intell., 29:2143–2156, 2007.

[84] Novi Quadrianto and Christoph H. Lampert. Learning multi-view neighborhood preserving projections. In ICML, 2011.

[85] Lishan Qiao, Songcan Chen, and Xiaoyang Tan. Sparsity preserving projections with applications to face recognition. Pattern Recognition, 43:331–341, 2010.

[86] Limei Zhang, Lishan Qiao, and Songcan Chen. Graph-optimized locality preserving projections. Pattern Recognition, 43:1993–2002, 2010.

[87] Limei Zhang, Songcan Chen, and Lishan Qiao. Graph optimization for dimensionality reduction with sparsity constraints. Pattern Recognition, 45:1205–1210, 2012.

[88] Lishan Qiao, Limei Zhang, and Songcan Chen. Dimensionality reduction with adaptive graph. Frontiers of Computer Science, 7:745–753, 2013.

[89] Shuicheng Yan, Dong Xu, Benyu Zhang, HongJiang Zhang, Qiang Yang, and Stephen Lin. Graph embedding and extensions: A general framework for dimensionality reduction. IEEE Trans. Pattern Anal. Mach. Intell., 29:40–51, 2007.

[90] Haifeng Li, Tao Jiang, and Keshu Zhang. Efficient and robust feature extraction by maximum margin criterion. In IEEE Transactions on Neural Networks, 2003.

[91] Deng Cai, Xiaofei He, Kun Zhou, Jiawei Han, and Hujun Bao. Locality sensitive discriminant analysis. In IJCAI, 2007.

[92] Hwann-Tzong Chen, Huang-Wei Chang, and Tyng-Luh Liu. Local discriminant embedding and its variants. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), 2005.

[93] Jiani Hu, Weihong Deng, Jun Guo, and Weiran Xu. Learning a locality discriminating projection for classification. Knowl.-Based Syst., 22:562–568, 2009.

[94] Jian Cheng, Qingshan Liu, Hanqing Lu, and Yen-Wei Chen. Supervised kernel locality preserving projections for face recognition. Neurocomputing, 67:443–449, 2005.

[95] Wankou Yang, Changyin Sun, and Lei Zhang. A multi-manifold discriminant analysis method for image feature extraction. Pattern Recognition, 44:1649–1657, 2011.

[96] Jiwen Lu, Yap-Peng Tan, and Gang Wang. Discriminative multimanifold analysis for face recognition from a single training sample per person. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011.

[97] R. Hettiarachchi and James F. Peters. Multi-manifold lle learning in pattern recognition. Pattern Recognition, 48:2947–2960, 2015.

[98] Deng Cai, Xiaofei He, and Jiawei Han. Semi-supervised discriminant analysis. In 2007 IEEE 11th International Conference on Computer Vision, 2007.

[99] Yangqiu Song, Feiping Nie, and Changshui Zhang. Semi-supervised sub-manifold discriminant analysis. Pattern Recognition Letters, 29:1806–1813, 2008.

[100] Masashi Sugiyama, Tsuyoshi Idé, Shinichi Nakajima, and Jun Sese. Semi-supervised local fisher discriminant analysis for dimensionality reduction. In PAKDD, 2008.

[101] Andrew B. Goldberg, Xiaojin Zhu, Aarti Singh, Zhiting Xu, and Robert D. Nowak. Multi-manifold semi-supervised learning. In AISTATS, 2009.

[102] 苏煜, 山世光, 陈熙霖, and 高文. 基于全局和局部特征集成的人脸识别. 软件学报, 21(8):1849–1862, 2010.

[103] Rajkiran Gottumukkal and Vijayan K. Asari. An improved face recognition technique based on modular pca approach. Pattern Recognition Letters, 25:429–436, 2004.

[104] Songcan Chen and Yulian Zhu. Subpattern-based principal component analysis. 2004.

[105] Keren Tan and Songcan Chen. Adaptively weighted sub-pattern pca for face recognition. Neurocomputing, 64:505–511, 2005.

[106] Kadappagari Vijaya Kumar and Atul Negi. Subxpca and a generalized feature partitioning approach to principal component analysis. Pattern Recognition, 41:1398–1409, 2008.

[107] Timo Ahonen, Abdenour Hadid, and Matti Pietikäinen. Face recognition with local binary patterns. In ECCV, 2004.

[108] Chengjun Liu and Harry Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Trans. Image Processing, 11:467–476, 2002.

[109] Wenchao Zhang, Shiguang Shan, Wen Gao, Xilin Chen, and Hongming Zhang. Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, 2005.

[110] Loris Nanni and Dario Maio. Weighted sub-gabor for face recognition. Pattern Recognition Letters, 28:487–492, 2007.

[111] Stan Z. Li, XinWen Hou, HongJiang Zhang, and QianSheng Cheng. Learning spatially localized, parts-based representation. In CVPR, 2001.

[112] Yu-Lian Zhu. Sub-pattern non-negative matrix factorization based on random subspace for face recognition. In Wavelet Analysis and Pattern Recognition, 2007. ICWAPR'07. International Conference on, volume 3, pages 1356–1360. IEEE, 2007.

[113] David L Donoho. Compressed sensing. IEEE Transactions on information theory, 52(4):1289–1306, 2006.

[114] John Wright, Allen Y Yang, Arvind Ganesh, S Shankar Sastry, and Yi Ma. Robust face recognition via sparse representation. IEEE transactions on pattern analysis and machine intelligence, 31(2):210–227, 2009.

[115] Weifeng Liu, Dacheng Tao, Jun Cheng, and Yuanyan Tang. Multiview hessian discriminative sparse coding for image annotation. Computer Vision and Image Understanding, 118:50–60, 2014.

[116] Stéphane Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. IEEE Trans. Signal Processing, 41:3397–3415, 1993.

[117] Y. C. PATI and R. REZAIIFAR. Orthogonal matching pursuit: Recursive function approximat ion with applications to wavelet decomposition. 1993.

[118] Scott Saobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. SIAM Review, 43:129–159, 1998.

[119] Duc-Son Pham and Svetha Venkatesh. Joint learning and dictionary construction for pattern recognition. In 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008.

[120] Qiang Zhang and Baoxin Li. Discriminative k-svd for dictionary learning in face recognition. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010.

[121] Zhuolin Jiang, Zhe L. Lin, and Larry S. Davis. Learning a discriminative dictionary for sparse coding via label consistent k-svd. In CVPR, 2011.

[122] Julien Mairal, Francis R. Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. Discriminative learned dictionaries for local image analysis. In 2008 IEEE Conference on Computer Vision and Pattern Recognition, 2008.

[123] Meng Yang, Lei Zhang, Xiangchu Feng, and David Zhang. Fisher discrimination dictionary learning for sparse representation. In 2011 International Conference on Computer Vision, 2011.

[124] Xiao-Chen Lian, Zhiwei Li, Bao-Liang Lu, and Lei Zhang. Max-margin dictionary learning for multiclass image categorization. In ECCV, 2010.

[125] Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. 2004.

[126] Rodolphe Jenatton, Julien Mairal, Guillaume Obozinski, and Francis R. Bach. Proximal methods for sparse hierarchical dictionary learning. In ICML, 2010.

[127] Yangqing Jia, Mathieu Salzmann, and Trevor Darrell. Factorized latent spaces with structured sparsity. In NIPS, 2010.

[128] Haichao Zhang, Nasser M. Nasrabadi, Yanning Zhang, and Thomas S. Huang. Joint dynamic sparse representation for multi-view face recognition. Pattern Recognition, 45:1290–1298, 2012.

[129] Miao Zheng, Jiajun Bu, Chun Chen, Can Wang, Lijun Zhang, Guang Qiu, and Deng Cai. Graph regularized sparse coding for image representation. IEEE Trans. Image Processing, 20:1327–1336, 2011.

[130] Miao Zheng, Jiajun Bu, and Chun Chen. Hessian sparse coding. Neurocomputing, 123:247–254, 2014.

[131] Shenghua Gao, Ivor W. Tsang, and Liang-Tien Chia. Laplacian sparse coding, hypergraph laplacian sparse coding, and applications. IEEE Trans. Pattern Anal. Mach. Intell., 35:92–104, 2013.

[132] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1701–1708, 2014.

[133] Gary B Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A sdatabase for studying face recognition in unconstrained environments. Technical report, Technical Report 07-49, University of Massachusetts, Amherst, 2007.

[134] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation from predicting 10,000 classes. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1891–1898, 2014.

[135] Yi Sun, Yuheng Chen, Xiaogang Wang, and Xiaoou Tang. Deep learning face representation by joint identification-verification. In Advances in neural information processing systems, pages 1988–1996, 2014.

[136] Yi Sun, Xiaogang Wang, and Xiaoou Tang. Deeply learned face representations are sparse, selective, and robust. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2892–2900, 2015.

[137] Yi Sun, Ding Liang, Xiaogang Wang, and Xiaoou Tang. Deepid3: Face recognition with very deep neural networks. arXiv preprint arXiv:1502.00873, 2015.

[138] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556, 2014.

[139] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. CoRR, abs/1409.4842, 2015.

[140] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 815–823, 2015.

[141] Jingtuo Liu, Yafeng Deng, Tao Bai, and Chang Huang. Targeting ultimate accuracy: Face recognition via deep embedding. CoRR, abs/1506.07310, 2015.

[142] Bilei Zhu and Hong Liu. Mirex 2015 qbsh task: Tencent bestimage's solution. 2015.

[143] Erjin Zhou, Zhimin Cao, and Qi Yin. Naive-deep face recognition: Touching the limit of lfw benchmark or not? CoRR, abs/1501.04690, 2015.

[144] Brendan Klare, Benjamin Klein, Emma Taborsky, Austin Blanton, Jordan Cheney, Kristen Allen, Patrick Grother, Alan Mah, Mark Burge, and Anil K. Jain. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[145] Ira Kemelmacher-Shlizerman, Steven M. Seitz, Daniel Miller, and Evan Brossard. The megaface benchmark: 1 million faces for recognition at scale. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[146] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In ECCV, 2016.

[147] Iacopo Masi, Stephen Rawls, Gérard G. Medioni, and Premkumar Natarajan. Pose-aware face recognition in the wild. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[148] Meina Kan, Shiguang Shan, and Xilin Chen. Multi-view deep network for cross-view classification. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[149] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron C. Courville, and Yoshua Bengio. Generative adversarial nets. In NIPS, 2014.

[150] Di He, Yingce Xia, Tao Qin, Liwei Wang, Nenghai Yu, Tie-Yan Liu, and Wei-Ying Ma. Dual learning for machine translation. In NIPS, 2016.

[151] 张钹. 人工智能未来展望，后深度学习时代. In 2016 China National Computer Congress (CNCC), 2016.